Estimation of the dynamics and yields of cereals in a semi-arid area using remote sensing and the SAFY growth model

Aicha Chahbi, Mehrez Zribi, Zohra Lili-Chabaane, Benoît Duchemin, Marouen Shabou, Bernard Mougenot, Gilles Boulet

Estimation of the dynamics and yields of cereals in a semi-arid area using remote sensing and the SAFY growth model

Aicha Chahbi[1-2], Mehrez Zribi[1], Zohra Lili-Chabaane[2], Benoit Duchemin, Marouen Shabou[1-2], Bernard Mougenot[1], Gille Boulet[1]

1-CESBIO, 18 avenue. Edouard Belin, bpi 2801, 31401 Toulouse cedex 9, France

2- INAT-LRSTE (Université de Carthage), 43, Avenue Charles Nicolle 1082 -Tunis-Mahrajène, Tunisia

ABSTRACT

In semi-arid areas, high climatic variability induce a major risk on food security. Therefore an operational forecasting system for the grain yields is required and could help decision-makers to make early decisions and plan annual imports. It can be challenging to monitor the crop canopy and production capacity of plants, especially cereals. In this context, the aim of the present study is to analyze the characteristics of two types of irrigated and non-irrigated cereals: barley and wheat. Through the use of a rich database, acquired over a period of two years for more than 30 test fields, and from 20 optical satellite SPOT/HRV images, two research approaches are considered. Firstly, statistical analysis is used to characterize the vegetation's dynamics and grain yield, based on remotely sensed (satellite) Normalized Difference Vegetation Index (NDVI) measurements. A relationship is established between the NDVI and LAI (leaf area index). Different robust relationships (exponential or linear) are established between the satellite NDVI index acquired from SPOT/HRV images, just before the time of maximum growth (April), and grain and straw, for barley and wheat vegetation covers. Following validation of the proposed empirical approaches, yield maps are produced

for the studied site. The second approach is based on the application of a SAFY "Simple Algorithm For Yield estimation" growth model, developed to simulate the dynamics of the LAI and the grain yield. An inter-comparison between ground yield measurements and SAFY model simulations reveals that the yields are under-estimated by this model. Finally, the combination of multi-temporal satellite measurements with the SAFY model estimations is also proposed for the purposes of yield mapping. The SAFY results are found to be reasonably correlated with those determined by satellite measurements (NDVI). Nevertheless, it should be noted that the SAFY model still underestimates the yields.

1. INTRODUCTION

The world's population has been constantly increasing over the last half century. This growth has been accompanied by an increase in the consumption of agricultural products, especially cereals, such that it is becoming increasingly important for scientists and decision-makers to be able to estimate grain yields at regional and national levels. An operational grain-yield forecasting system is thus required, and could help decision-makers to make timely decisions and plan annual imports (Justice and Becker-Reshef, 2007). Many models, based on the use of remote sensing or agro-meteorological models, have been developed to estimate the biomass and grain yield of cereals (Barnett and Thompson 1982, Moriondo *et al.* 2007, Balaghi *et al.* 2008, Laurila *et al.* 2010, Benedetti and Rossinni 1993, Becker-Reshef *et al.* 2010).

Remote sensing has demonstrated its strong potential for the monitoring of the soil and vegetation's dynamics and temporal variations, mainly because it provides wide spatial coverage and has consistent internal datasets (Zribi et al., 2008, Amri et al., 2011). Optical remote sensing has demonstrated its strong potential for the monitoring of the vegetation's dynamics and temporal variations, mainly because it provides wide spatial coverage and has consistent internal datasets. In particular, the Normalized Difference Vegetation Index (NDVI) expresses the contrast in reflectance between the red and near-infrared regions of a

surface spectrum (Rouse *et al.* 1974). The NDVI is easy to calculate, which can be related to the green vegetation cover or vegetation abundance, and is expressed by: NDVI = (NIR − RED)/(NIR + RED), where NIR is the near-infrared reflectance and RED is the red reflectance. The NDVI is sensitive to the presence of green vegetation (Sellers *et al.* 1985) and has been used for several regional and global applications, in studies concerning the distribution and potential photosynthetic activity of vegetation (Deblonde *et al.* 1993, Propastin *et al.* 2009). Efforts have been made to develop various models and approaches for the purposes of forecasting the yields of different crops, in different regions throughout the globe. Most studies related to the estimation of cereal yields make use of low resolution sensors (Balaghi *et al.* 2008, Kogan *et al.* 2013, Bastiaanssen and Ali 2003). Balaghi *et al.* (2008) propose the use of empirical ordinary least squares regression models to forecast yields at provincial and national levels. Their predictions are based on decadal NDVI/AVHRR data, decadal accumulated rainfall, and average monthly air temperatures. Becker-Becker-Reshef *et al.* (2010) combine a new BRDF-corrected daily surface reflectance dataset, developed from NASA's Moderate resolution Imaging Spectro-radiometer (MODIS), with detailed official crop statistics. They use this combined data to develop an empirical, generalized approach to the forecasting of wheat yields. This is based on a single, generalized model, which has been applied at the level of the state of Kansas (USA) and has been shown to be directly applicable to the Ukraine. Yield estimations are also relevant to other types of vegetation, including rice (Wang *et al.* 2010, Wei-guo *et al.* 2011), soybean and soya (Prasad *et al.* 2006), olives (Maselli *et al.* 2012), citrus fruit (Ye *et al.* 2006), maize and sunflower crops (Claverie *et al.* 2012).

In recent decades, there has also been a growing interest in approaches based on agro-ecological process models combined with remote sensing. Vegetation models based on climatic and agronomic data have demonstrated a good degree of correlation between the

estimated and measured variables, at the local level (Duchemain *et al.* 2008, Bastiaanssen and Ali 2003). These models simulate key variables, including the LAI (Leaf Area Index). The model proposed by Bastiaanssen and Ali (2003) combines the photosynthetically active radiation (PAR) model of Monteith (1972) with the light-use efficiency model of Field *et al.* (1995) and the surface energy balance model of Bastiaanssen *et al.* (1998), to estimate the growth of irrigated crops. To forecast crop yields, these authors use NOAA's AVHRR instrument (Advanced Very High Resolution Radiometer) during periods of crop irrigation. Most of these studies have been carried out in Europe or North America, where generally high levels of cereal production are encountered.

Lobell *et al.* (2003) estimated crop yields using AVHRR imagery over northwest Mexico. Remotely-sensed estimations of the fraction of absorbed photosynthetically active radiation (fAPAR) were incorporated into a simple model based on crop light-use efficiency estimations derived from Monteith (1972, 1977), to predict wheat yields and optimal planting dates.

The aim of the present study is to evaluate the feasibility of three approaches: high resolution remote sensing imagery, the agro-meteorological SAFY model, and combined use of the SAFY model with remotely sensed data, to estimate the dynamics and yields of cereals in the context of semi-arid, low productivity regions in North Africa.

Section 2 describes the studied site and the experimental satellite and ground measurement database. Section 3 is dedicated to the statistical analysis of this data, used to estimate the dynamics and yields of cereals, derived from the remotely sensed NDVI. In section 4, the use of the SAFY model, for the estimation of vegetation cycle dynamics and yields, is described. Then, the combined use of the SAFY model and remotely sensed data is presented. Our conclusions are given in section 5.

2. Experimental database

2.1 Study area

The Kairouan plain (Zribi *et al*. 2011) is situated in central Tunisia (9°30’ E, 10°15’E, 35°N, 35°45’ N) (Fig. 1). The climate in this region is semi-arid, with an average annual rainfall of approximately 300 mm per year, characterized by a rainy season lasting from October to May, with the two rainiest months being October and March. As is generally the case in semi-arid areas, the rainfall patterns in this area are highly variable in time and space. The mean temperature in Kairouan City is 19.2 °C (minimum of 10.7 °C in January and maximum of 28.6 °C in August) and the mean annual potential evapotranspiration (Penman) is close to 1600 mm. The landscape is mainly flat and the vegetation in this area is dominated by agriculture (cereals, olive trees, and market gardens). There are various types of crop, and their rotation is typical of semi-arid regions. The aquifer of the Kairouan plain represents the largest basin in central Tunisia. It is fed by the infiltration of surface waters during floods in the natural regime, or at the time of dam releases since the construction of the Sidi Saad and El Haouareb dams. Surface and groundwater streams drain into the Sebkha Kelbia, a large salt lake.

2.2 Satellite Data

For the purposes of monitoring surface parameters and gaining an understanding of the vegetation dynamics on the Kairouan plain, **15** images acquired by the SPOT 5 satellite were analyzed (Table1). These images were acquired with a repeat time of approximately 21 days and a high spatial resolution, equal to 10 meters. All of the acquired images were orthorectified using the Ground Control Point technique. The resulting mismatch between these images and the reference image was less than ± 0.5 pixel. In a second step, the orthorectified images were calibrated to derive the TOA (Top of Atmosphere) reflectances,

and then atmospherically corrected to obtain the TOC (Top of Canopy) reflectances. The raw digital values were initially converted into TOA reflectances using the same absolute calibration techniques as those which are operationally applied to SPOT image data (Meygret 2005). The atmospheric correction needed to determine the TOC reflectances relies on inversion of the 6S radiative transfer model (Berthelot and Dedieu 1997, Rahman and Dedieu 1994), based on the use of look-up tables. This model requires input parameters such as the aerosol optical thickness (AOT) at 550 nm, size distribution and refractive index of the aerosols, water vapor content, ozone content, and atmospheric pressure. The water vapor content was determined from measurements recorded at the studied site, the ozone content was obtained from the Ozone Monitoring Instrument (OMI) data which is available at http://toms/gsfc.nasa.gov, and the surface atmospheric pressure was retrieved from http://www.tutiempo.net.

2.3 Ground Measurements

Simultaneously to the satellite data acquisitions, ground campaigns were carried out approximately once every two weeks on a large number of test plots, during two agricultural seasons (2010-2011 and 2011-2012). During the first year a total of 27 test fields, including 18 irrigated and 9 non-irrigated cereal fields, were used for the ground truth measurements (Fig. 1), whereas during the second season, 55 test fields, including 36 irrigated and 19 non-irrigated cereal fields, were analyzed. Two types of cereal were considered: wheat and barley in the case of the selected test fields. Table 2 provides a detailed breakdown of the characteristics of these fields.

The *in situ* measurements involved mainly: water content, height, LAI (Leaf Area Index) of the vegetation, and cereal yields.

a) Leaf Area Index

The LAI is defined as the total one-sided area of leaf tissue per unit of ground surface area. According to this definition, the LAI is a dimensionless quantity characterizing the canopy of an ecosystem. Two different approaches can be used to estimate the LAI: directly, through destructive sampling, or indirectly, by applying methods based on ground-level gap fraction measurements. In the present study, an indirect method is used, in which the LAI is derived from hemispherical digital photography. A binary classification of green elements and soil is proposed, in order to compute the gap fraction at a 57.5° zenith angle, from which an estimation of the LAI is then derived (Kirk *et al.* 2009, Baret *et al.* 2010, Sandman *et al.* 2013).

These measurements were applied to each cereal field, on different days during the vegetation season. Thirty LAI estimations were made for each field, inside two 20 x 20 m^2 square areas. The LAI measurements were made during the same week as that during which the SPOT 5 satellite images were recorded, in order to limit any discrepancies between ground and satellite measurements.

Fig. 2-a plots the variations in these measured quantities for three test fields, two of which were irrigated. The third field was rainfed during the second season. Firstly, a LAI maximum, which occurred after a period of accelerated growth, can be observed during the first two weeks of April. Secondly, a clear difference can be seen between the irrigated and rainfed vegetation cycles, with a greater increase in the LAI parameter at the beginning of the vegetation cycle in the case of the irrigated fields.

b) Cereal yield

For each test field, ten sample measurements were taken along the two field diagonals, using a 75 x 75 cm^2 frame. The cereal yields were then measured: the number of stalks, the weight of the grain, and the weight of the straw were recorded. The yields and quantities of

grain lost during harvesting were studied simultaneously. As a result of the use of data taken from both rainfed and irrigated fields, the estimated yields have a large dynamic range, between 1300 kg/ha and 10300 kg/ha, with a mean value of 4300 kg/ha for wheat, and between 1500 kg/ha and 5500 kg/ha, with a mean value of 3400 kg/ha, for barley. Figure 2-b shows the error bars corresponding to the measured grain yield and the NDVI extracted from satellite images at the beginning of April, for the 82 test fields. The mean standard deviation of the NDVI is 0.067, with an average value for all the test fields equal to 0.58. For the measured grain yields, the average value for all of the fields is equal to 3400 kg/ha, and the mean standard deviation for all of the test fields is equal to 970 kg/ha.

2.4 Land use mapping

Land use identification (Fig. 3) is an essential step in the mapping of yields over all cereal fields. For this, land use mapping was implemented, based on a decision tree and using two types of satellite data: four SPOT images acquired at four different times of the year were used to identify different vegetation cycles, and SRTM data was used for relief identification. A decision tree classification approach was used, from which eleven classes of land use were identified: pluvial olive trees, irrigated olive trees, irrigated winter vegetables, irrigated summer vegetables, bare soils, cereals, urban areas, mountainous areas, water cover (wadi beds and dams), and coastal salt flats ("sebkhas"). For all of these classes, thresholds were considered for one or two images. Test fields were used for the learning phase to identify empirical thresholds of NDVI in order to separate the different classes. For example, for summer vegetables, an empirical NDVI threshold was applied to the image acquired in July (NDVI>0.3). In the case of winter vegetables, the image acquired at the end of autumn or the beginning of summer was used. During these periods of the year, vegetables have the highest NDVI index. For the cereal classes (irrigated or non-irrigated, wheat or barley), the satellite data acquired in March or April, corresponding to the cereals' maximum growth, was used.

The DTM provided by the Shuttle Radar Topography Mission (SRTM, http://srtm.usgs.gov/) allowed high relief areas to be eliminated from the land use analysis. In order to validate these remotely sensed classifications, a confusion matrix was produced by comparing the classification results using more than 100 test fields, characterized by different types of land use. This analysis has revealed an overall accuracy of approximately 80%, and accuracy greater than 90% in the case of cereals. The non-irrigated olive tree class covers 43% of the studied site, whereas the wheat class corresponds to 12% of the surface area of the studied site.

Once established, the land use map can be used to extract a cereal mask for the two agricultural seasons under study.

3. Statistical analysis of the relationship between cereal yields and satellite observations

The aim of this section is to evaluate the potential of remote sensing, combined with statistical analysis, for the estimation of barley and wheat yields.

3.1 Relationship between SPOT-NDVI and cereal yields

Fig. 4 illustrates the relationship between the NDVI, acquired on 17/03/2011 and 31/03/2012, and the grain and straw yield measurements (Y_{grain} and Y_{straw}) obtained at the end of the cereal season. For the first dataset, including 44 fields and covering a large range of sowing dates, as well as irrigation and fertilization schedules, two exponential relationships were determined empirically:

$$Y_{grain} = 4.34 \exp(3.38 \times NDVI)$$
$$Y_{straws} = 2.85 \exp(3.88 \times NDVI) \quad (1)$$

The NDVI is well correlated with the average weight of the harvested grain (R^2=0.66), and with the average weight of the straw (R^2 = 0.74). In the case of the data acquired before and

after that corresponding to Fig. 4, the correlation between the satellite index and the measured yields is weaker. In practice, the end of March corresponds to a maximum in the vegetation dynamics. Before this date, NDVI variations between fields are not affected by irrigation or rainfall events occurring in March, which are essential to grain development. In mid-April, the vegetation LAI starts to decrease, leading to a reduced capacity to distinguish between high and low grain yields.

Fig. 5 plots the NDVI-estimated grain and straw yields as a function of the ground truth measurements, thus validating the proposed empirical algorithms for the second set of test fields: the resulting RMSE is equal to 850 kg/ha for the grain yields and 1160 kg/ha for the straw yields. This outcome demonstrates the robustness of the proposed empirical approach, despite its simplicity.

The same yield estimation algorithm was applied separately for the wheat and barley crops, independently of the various agricultural practices (sowing date, type of seeder, irrigation date, quantity of fertilizer, …). The results are described below.

a) Wheat yields

Fig. 6 shows a plot of the exponential relationship between the measured yields (grain and straw) and the NDVI, for the first set of tested wheat fields. The measured yield and the NDVI can be seen to be well correlated, with correlation coefficients R^2 equal to 0.67 and 0.66, for grain and straw respectively.

The empirical relationships were validated with a second set of test fields. The measured and estimated yields can be seen to be in good agreement, with an RMS error equal to 900 kg/ha and 1100 kg/ha, for grain and straw respectively (Fig. 7).

b) Barley yields

Twenty-two barley fields were observed during the two agricultural years. The database was divided into two sets: the first was used to estimate the yield on the basis of the NDVI measurements, and the second set was used for model validation. Fig. 8 provides a plot of the linear relationship established between measured and NDVI-derived grain and straw yields, for the first set of test fields. The correlation coefficients R^2 are equal to 0.62 and 0.70, for grain and straw respectively.

The proposed relationship was validated by comparing yield measurements and estimations with the second set of fields. The results are found to be in good agreement, with an RMS error equal to 650 kg/ha and 1470 kg/ha, for the grain and straw estimations respectively (Fig. 9). In the case of barley, only a linear relationship was considered, since the number of test fields was too small for a more complex (e.g. exponential) relationship to be determined.

3.2 Yield mapping

Following the above-described validation, Eq. 1 was used to invert the NDVI and generate yield maps, for all classes of cereal. The wheat and barley classes could not be treated separately, because they were not differentiated by the land use identification process.

In order to eliminate the influence of local terrain heterogeneities (due to local variations in vegetation cover and soil surface, etc.) on the processed optical signal, the yield was estimated over cells corresponding to 10 x 10 pixels (approximately 100 m^2), rather than at the scale of single pixels. For each resulting cell, the yield estimation was applied only if more than 25% of the cell's pixels belonged to cereal fields. The value of the computed yield could then be considered as representative of the entire cell. In Fig. 10 the yield maps are shown for the 2011-2012 agricultural season.

4. Estimation of the dynamics and yields of cereals using the SAFY growth model

4.1 SAFY Model description

One of the aims of the present study was to evaluate the potential use of times series of a variable biophysical parameter (the LAI) for the monitoring of phytomass production and the grain yields of cereal crops. This evaluation was based on the use of the ‘‘Simple Algorithm For Yield estimation’’ (SAFY) model (Duchemin *et al*. 2008). The main idea was to use the model to represent well-known processes involved in crop development and growth, with the requirement that these processes be simulated using standard data, i.e. climatic data and optical imagery (which provides LAI estimations). Climatic data were generated by a weather station located in our study area (lat: 35.56°, long: 9.94°), which provides air temperature (Ta) and incoming global radiation (Rg) measurements. This data is recorded once every 30 minutes.

The model simulates the increase in dry above-ground phytomass, based on the light-use efficiency theory of Monteith (1977), and also takes the dynamics of green leaves and the influence of temperature into account (Duchemin *et al*, 2008). The dynamics of the green LAI are simulated by estimating the extent of leaf coverage during growth (ΔLAI+) and the decrease in LAI during senescence due to the wilting and falling of leaves (ΔLAI-). These two phenological phases are identified through the use of a degree-day approach based on the accumulated air temperature (ΣT_a). Leaf senescence starts when ΣT_a has attained a given threshold S_{TT} (sum of temperature for senescence).

$$\begin{aligned} \Delta LAI+ &= \Delta DAM \times P_L\left(\Sigma T_a\right) \times SLA \\ \Delta LAI- &= LAI\left(\Sigma T_a - S_{TT}\right) / R_s \end{aligned} \qquad (2)$$

Where DAM is the dry, above-ground mass, P_L is the partition-to-leaf function, which can be expressed as an empirical function of the air temperature, SLA is the specific leaf area, and R_s is the rate of senescence.

SAFY has a low level of complexity, in order to simplify the optimization of unknown parameters when a small number of observations is available. The parameters are limited in number (14), and are used to determine *a priori* values based on data provided by various prior experimental studies (Claverie *et al.* 2012, Meinke *et al*. 1998, Porter and Gawith 1999, Sinclair and Amir 1992, Varlet-Grancher *et al*. 1982). However, three of these parameters (the day of plant emergence D_0, the effective light-use efficiency ELUE and the "sum of temperature for senescence" S_{TT}) are strongly dependent on the ambient agro-environmental conditions.

4.2 Application of the SAFY model to cereal cycle retrieval

The first step involves calibrating the three parameters (D_0, ELUE, S_{TT}). This procedure corresponds to the identification of an optimum parameter set from which, for each field, the SAFY simulation leads to the best reproduction of several observed variables. These parameters were calibrated using an optimization algorithm based on the observed ground truth LAI values collected in the fields. For each test plot, three optimized parameters were determined. These provided the “best combination”, defined as that having the lowest Root Mean Square Error (RMSE) between the LAI measured in the field and that estimated by the SAFY model. Table 4 indicates the range over which each of these parameters was considered.

The values of these three parameters are thus determined, for each test plot, on the basis of LAI measurements. The cereal dynamics are derived, for each field, from the estimated LAI. In Fig. 11, the LAI determined with the model is compared with that measured on the ground, for four different fields planted with irrigated and rainfed cereals. It can be seen that in all four cases, the model accurately retrieves the dynamic range of the LAI. Maximum growth occurs at the end of March.

4.3 Application of the SAFY model to yield estimations

The grain-filling phase ranges from the day when foliage production ends, to the day when total senescence occurs. During this period, the daily increase in grain yield (ΔGY) is proportional to the total above-ground phytomass (DAM), with a constant fraction P_y partitioned to grains. This leads to:

$$\Delta GY = P_y \times DAM \quad (3)$$

Since the characteristic parameters of the wheat and **barley** fields are quite different, the SAFY model was validated independently for each type of cereal. Field measurements were thus used to calibrate the fraction Py, for each type of cereal, by minimizing the RMSE between the grain yields measured in all of the fields, and the wheat and barley grain yields simulated by the SAFY model (Py_{wheat} = 0.0108 and Py_{barley} = 0.0057).

Fig. 12-a compares the SAFY model yield estimations with ground measurements, for the case of the barley fields. Although the ground measurements and model estimations are only moderately correlated **(R^2_{barley}=0.35),** the SAFY model can be seen to produce results which are coherent with the ground measurements in terms of yield dynamics. The limitations of the SAFY model can be explained by various influences (fertilization, irrigation…), which it does not take into account. In Fig. 12-b, the SAFY model wheat yield estimations are compared with the ground measurements. As in the case of the barely fields, only a moderate degree of correlation is found (R^2_{wheat}=0.33) between the two types of data.

4.4 Application combining the SAFY model with remotely sensed data

In this section, multi-temporal satellite LAI measurements are used to calibrate the SAFY model. Before implementing this calibration, a relationship is established between the satellite NDVI and the LAI.

a) NDVI-LAI relationship for cereals

The LAI variable can be determined **from the** NDVI. Several types of relationship can be found in the literature, comparing these two key parameters. According to Wardley and Curran (1984), the NDVI-LAI relationship can be assumed to be linear, although when the LAI lies in the interval between 2 and 6 this relationship is no longer valid. Asrar *et al*. (1984) and Richardson *et al*. (1992) derived logarithmic relationships to express the NDVI as a function of the LAI.

In the present study, for each pixel a NDVI profile was generated from the satellite data. A relationship was then established between the NDVI and LAI of the studied test fields, during the growing seasons of 2008-2009 and 2010-2011. As proposed by Duchemin *et al*. (2006), an exponential relationship of the following type was then determined:

$$NDVI = NDVI_{\infty} + \left(NDVI_{soil} - NDVI_{\infty}\right) \times e^{-k_{NDVI} \; LAI} \qquad (4)$$

- where $NDVI_{\infty}$ is the asymptotic value of the NDVI, derived from the measurements, In the case of the present study, $NDVI_{\infty}$ was taken to be 0.75 (see Fig. 13-a), and

- $NDVI_{soil}$ is the value of the NDVI for bare soil (value determined in the present study: 0.15),

- k_{NDVI} is the extinction coefficient.

Fig. 13-a plots the experimental LAI – NDVI data points, corresponding to the 2009-2010 and 2011-2012 seasons, together with a least-squares fit to the above empirical relationship, which has a correlation coefficient equal to 0.78. It should be noted that when the LAI exceeds 2, the NDVI saturates (at approximately 0.75), leading to NDVI errors at high LAI values.

The proposed algorithm was validated by comparing ground-truth measurements with estimations derived from SPOT NDVI data for the test fields.

The resulting RMSE is equal to 0.73, as shown in Fig. 13-b. In accordance with the NDVI saturation effect noted above, the proposed empirical algorithm is found to become inaccurate for high LAI values.

b) Application of the SAFY model using satellite LAI estimations

In this section, the grain yield is estimated using the SAFY model combined with remotely sensed, multi-temporal SPOT acquisitions. This process involves initial calibration of the SAFY model, achieved by minimizing the RMSE between the SPOT5- time series LAI, and the SAFY model LAI. Maps of the three parameters: ELUE, D_0 and S_{TT}, optimized for the SAFY model, are thus generated. In order to eliminate the effects of local heterogeneities (due to irregularities occuring during the sowing and growth stages, vegetation dispersion heterogeneities, etc.), these 3 parameters are estimated over cells corresponding to 10 x 10 pixels (approximately 100 m^2). When more than 25% of each of cell belongs to the relevant class of cereals, the estimated values of the 3 corresponding parameters are applied.

After having generating the (D_0, ELUE, S_{TT}) parameter maps, an empirical relationship $\Delta GY = P_y \times DAM + b$ was used to generate (wheat and barley) cereal yield maps. The SAFY model thus estimates the grain yields over a 10 x 10 pixel window. The yield map corresponding to the 2011/2012 agricultural year is shown in Fig. 14.

Fig. 15-a compares, for all pixels, the yield estimations computed using statistical remote sensing analysis (section 3) with those determined by combining the SAFY model with remotely sensed measurements and in fig. 15-b we illustrate the difference, for each pixel, between the two maps of yields. The two estimations can be seen to be reasonably correlated, with R^2 equal to 0.45. Nevertheless, it should be noted that the SAFY model grain yields are underestimated. This could be explained by the fact that wheat and barley were not analyzed separately in this study. Also, the underestimation of LAI, due to the saturation effect of

NDVI (at approximately 0.75), could be one of reasons of the underestimation of the grain yield modeled by SAFY. In addition, the semi-empirical SAFY model has 14 input parameters, of which D_o, ELUE and S_{TT} were calibrated in the present study. However, there are three crop-specific parameters (Pla, Plb, Rs), the values of which depend on the type of cereal (wheat or barley). Another limitation of the SAFY model is that it assumes ELUE to remain constant over each cereal's phenological cycle, whereas other studies show that ELUE decreases during the maturity phase (Lecoeur *et al.* 2011, Claverie 2012).

5. CONCLUSION

The aim of this study was to characterize the dynamics and yields of various cereals in the semi-arid Merguellil catchment, using high-resolution optical data and a simple agro model.

In a first step, a statistical comparison between SPOT optical measurements and ground truth yields is proposed. The yield analysis for grain and straw as a function of NDVI reveals a strong correlation between these two variables, from mid-March to mid-April. An exponential relationship is retrieved between the yields and the remotely sensed NDVI index, with a correlation coefficient greater than 0.6. Validation of the remotely sensed estimations through the use of ground measurements shows that this approach is robust, with an rms error equal to 8.5 and 1160 kg/ha, for grain and straw yields, respectively. Exponential and linear relationships are determined for wheat and barley yields, based on the NDVI index. Validation of the proposed expressions leads to reasonable results, with an rms error lower than 900 kg/ha for both types of cereal. On the basis of the aforementioned analysis and the validation process, yield maps are proposed for all classes of cereal on the studied site.

In order to characterize the cereal cycle, a SAFY growth model is also used. This model provides excellent simulations of variations in the leaf area index, for irrigated and rainfed cereal fields. In the case of the yield estimations, despite the limited accuracy (low correlation

coefficient) of the model, coherent trends in the yield dynamics are retrieved for the studied test fields. The tested model has the advantage of being relatively simple, without requiring the input of data related to agricultural practices (sowing, irrigation and fertilization), and is thus very useful for operational applications on a regional scale.

Finally, the SAFY model is used in combination with remotely sensed LAI measurements. In an initial step, an exponential relationship is established between the NDVI and the LAI, both of which are derived from optical satellite data. An rms error equal to 0.78 is found when these remotely sensed parameters are validated through the use of ground measurements. This relationship allows the LAI to be mapped on cereal fields. Following calibration of the SAFY model through the use of multi-temporal LAI satellite maps, yield estimations are proposed for the entire studied site. The maps produced from satellite measurements are found to be reasonably correlated with those determined by combining satellite measurements and SAFY model data (R^2=0.45). The cereal yields predicted by the SAFY model are found to underestimate the real values measured on the ground.

ACKNOWLEDGMENTS:

This study was funded by two programs: the French SICMED/Mistrals program, and the French ANR's AMETHYST project. The authors extend their thanks to the officials of the Chebika Cereal Institute, the National Institute of Meteorology, and the Department of Soil and CRDA Kairouan, for their assistance during the test field measurements. The authors would also like to thank the technical team from the IRD (Institut de Recherche pour le Développement) and the INAT (Institut National Agronomique de Tunisie), for their collaboration in the acquisition of field data.

REFERENCES

Amri, R., Zribi, M., Duchemin, B., Lili-Chabaane, Z., Gruhier, C., Chebouni, A., Analysis of vegetation behaviour in a semi-arid region, using SPOT-VEGETATION NDVI data, Remote Sensing , 2011, *3*, 2568-2590.

Asrar, G., Fuchs, M., Kanemasu, E.T., Hatfield, J.L., 1984, Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. Agronomy Journal, **76**, pp. 300-306.

Balaghi, R., Tychon, B., Eerens, H., Jlibene, M., 2008, Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. International Journal of Applied Earth Observation and Geoinformation, **10**, pp. 438–452.

Baret, F., De Solan, B., Lopez-Lozano, R., Ma, K., Weiss, M., 2010, GAI estimates of row crops from downward looking digital photos taken perpendicular to rows at 57.5 degrees zenith angle: Theoretical considerations based on 3D architecture models and application to wheat crops. Agricultural and Forest Meteorology, **150**, pp. 1393-1401.

Barnett, T.L. and Thompson, D.R., 1982, The use of large-area spectral data in wheat yield estimation. Remote Sensing Of Environment, **12**, pp. 509-518.

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. J. Hydr. pp. 212–213, pp. 198-212.

Bastiaanssen, W.G.M. and Ali, S., 2003, A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. Agriculture, Ecosystems and Environment, **94**, pp. 321–340.

Becker-Reshef, I., Vermote, E., Lindeman, M., Justice C, 2010, A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sensing of Environment, **114**, pp. 1312–1323.

Benedetti, R. and Rossinni, P., 1993, On the use of NDVI profiles as a tool for agricultural statistics: the case study of wheat yield estimate and forecast in Emilia Romagna. Remote Sensing of Environment, **45**, pp. 311– 326.

Berthelot, B. and Dedieu, G., 1997, Correction of atmospheric effects for VEGETATION data. In :G.a.P. Guyot, T. (Editor), Physical Measurements and Signatures in Remote Sensing, Courchevel,France, pp. 19-25.

Claverie, M., Demarez, V., Duchemin, B., Hagolle, O., Ducrot, D., Marais-Sicre, C., Dejoux, J.F., Huc, M., Keravec, P., Béziat, P., Fieuzal, R., Ceschia, E., Dedieu, G., 2012, Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. Remote Sensing of Environment, 124, pp. 844-857.

Deblonde, G. and Cihlar, J., 1993, A multiyear analysis of the relationship between surface environmental variables and NDVI over the Canadian landmass. Remote Sensing Rev., **7**, pp. 151-177.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Ezzahar, J., Hoedjes, J.C.B., Kharrou, M.H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.C., Simmonneaux, V., 2006, Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crop coefficients, leaf area index and remotely-sensed vegetation indices. Agricultural Water Management,**79**, pp. 1-27.

Duchemin, B., Maisongrande, P., Boulet, G., Benhadj, I., 2008, A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index. Environmental Modelling & Software, **23**, pp. 876-892.

Field, C.B., Randerson, J.T., Malmstrom, C.M., 1995, Global net primary production: combining ecology and remote sensing. Remote Sensing Environment, **51**, pp. 74-88.

Justice, C. O. and Becker-Reshef, I., 2007, Report from the workshop on developing a strategy for global agricultural monitoring in the framework of Group on Earth Observations (GEO), pp. 1−67.

Kirk, K., Andersen, H.J., Thomsen, A.G., Jørgensen, J.R., Jørgensen, R.N., 2009, Estimation of leaf area index in cereal crops using red–green images. Biosystems Engineering, **104**, pp. 308- 317.

Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, O., Shelestov, A., Kolotii, A., Kussul, O., Lavrenyuk, A., 2013, Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. International Journal of Applied Earth Observation and Geoinformation, **23**, pp. 192-203.

Laurila, H., Karjalainen, M., Kleemola, J., Hyyppa, J., 2010, Cereal yield modeling in Finland using optical and radar remote sensing. Remote sensing, **2**, pp. 2185-2239.

Lecoeur, J., Poire-Lassus, R., Christophe, A., Pallas, B., Casadebaig, P., Debaeke, P., Vear, F., & Guilioni, L., 2011, Quantifying physiological determinants of genetic variation for yield potential in sunflower. SUNFLO: a model-based analysis. Functional Plant Biology, **38**, pp. 246-259.

Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L., 2003, Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agriculture, Ecosystems and Environment, **94**, pp. 205-220.

Maselli, F., Chiesi, M., Brilli, L., Moriondo, M., 2012, Simulation of olive fruit yield in Tuscany through the integration of remote sensing and ground data, Ecological Modelling, **244**, pp. 1-12.

Meinke, H., Hammer, G.L., van Keulen, H., Rabbinge, R., 1998, Improving wheat simulation capabilities in Australia from a cropping systems perspective III. The integrated wheat model (I_WHEAT). European Journal of Agronomy **8**, pp.101–116.

Meygret, A., 2005, Absolute Calibration: from SPOT1 to SPOT5, SPIE, San Diego.

Monteith, J.L., 1972, Solar radiation and productivity in tropical ecosystems. Journal of Applied Ecology, **9**, pp. 747-766.

Monteith, J.L., 1977, Climate and the efficiency of crop production in Britain. Philosphical Transaction of the Royal Society of London Ser. B, **281**, pp. 277-294.

Moriondo, M., Maselli, F., Bindi, M., 2007, A simple model of regional wheat yield based on NDVI data. European Journal of Agronomy, **26**, pp. 266-274.

Porter, J.R., Gawith, M., 1999, Temperatures and the growth and development of wheat: a review. European Journal of Agronomy **10**, pp. 23-36.

Prasad, A. K., Cha, L., Sing, R. P., Kafatos, M., 2006, Crop yield estimation model for Iowa using remote sensing and surface parameters. International Journal of Applied Earth Observation and Geoinformation, **8**, pp. 26-33.

Propastin, P. and Kappas, M., 2009, Modeling net ecosystem exchange for grassland in Central Kazakhstan by combining remote sensing and field data. Remote Sensing, **1**, pp. 159-183.

Rahman, H. and Dedieu, G., 1994, SMAC: A Simplified Method for the Atmospheric Correction of Satellite Measurements in the Solar Spectrum. International Journal of Remote Sensing, **15**, pp. 123-143.

Richardson, A.J., Wiegand, C.L., Wanjura, D.F., Dusek, D., Steiner, J.L., 1992, Multisite analyses of spectral biophysical data for sorghum. Remote Sensing of Environment, **41**, pp. 71-82.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974, Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. In Progress Report RSC 1978-1; Remote Sensing Center, Texas A&M University: College Station, TX, USA.

Sandmann, M., Graefe, J., Feller, C., 2013, Optical methods for the non-destructive estimation of leaf area index in kohlrabi and lettuce, Scientia Horticulturae, **156**, pp. 113-120.

Sellers, P.J., 1985, Canopy reflectance, photosynthesis and transpiration. International Journal Remote Sensing, **6**, pp. 1335-1372.

Sinclair, T.R., Amir, J., 1992, A model to assess nitrogen limitations on the growth and yield of spring wheat. Field Crops Research, **30**, pp. 63-78.

Varlet-Grancher, C., Bonhomme, R., Chartier, M., Artis, P., 1982, Efficience de la conversion de l'e´nergie solaire par un couvert végétal. Acta Oecologia/ Oecologia Plantarum, **17**, pp. 3–26.

Wang, Y.P., Chang, K.W., Chen, R.K., Lo, J.C., Shen, Y., 2010, Large-area rice yield forecasting using satellite imageries. International Journal of Applied Earth Observation and Geoinformation, **12**, pp. 27-35.

Wardley, N.W. and Curran, P.J., 1984, The estimation of green leaf area index from remotely sensed airborne multispectral scanner data. International Journal of Remote Sensing, **4**, pp. 671-679.

Wei-guo, L., Hua, L., Li-Hua, Z., 2011, Estimating rice yield by HJ-1A satellite images. Rice Science, **18(2)**, pp. 142-147.

Ye, X., Sakai, K., Garciano, L.O.,1, Asada, S.I., Sasao, A., 2006, Estimation of citrus yield from airborne hyperspectral images using a neural network model. Ecological modeling, **198**, pp. 426–432.

Zribi, M., Chahbi, A., Shabou, M., Lili-Chabaane, Z., Duchemin, B., Baghdadi, N., Amri, R., Chehbouni, G., 2011, Soil surface moisture estimation over a semi-arid region using ENVISAT ASAR radar data for soil evaporation evaluation. Hydrology Earth System Sciences, **15**, pp. 345-358.

Zribi, M., André, C., Decharme, B., 2008, A method for soil moisture estimation in Western Africa based on ERS Scatter meter, IEEE Transactions on Geoscience and Remote Sensing**,** 46, 2, 438-448.

FIGURE AND TABLE CAPTIONS

Figure 1. Illustration of the studied site.

Figure 2-a. Illustration of the evolution of LAI measurements for three test fields: irrigated (P14 and P9) and rainfed (P17) fields, during the 2011/2012 agricultural season.

Figure 2-b. Comparisons between the error bars of measured grain yield and the NDVI extracted from satellite images for barley (a) and wheat (b).

Figure 3. Land use map for the 2011-2012 agricultural season.

Figure 4. Relationship between mesured cereal yields (grain and straw) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012: (a) grain yields, (b) straw yields.

Figure 5: Comparison between measured and estimated cereal yields: (a) grain yields, b) straw yields.

Figure 6. Relationship between mesured wheat yields (grain (a) and straw (b)) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012.

Figure 7. Comparison between measured and estimated wheat yields: a) grain yields, b) straw yields.

Figure 8. Relationship between measured barley yields (grain (a) and straw (b)) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012.

Figure 9: Comparison between measured and estimated barley yields: a) grain yields, b) straw yields.

Figure 10. Cereal yield map for the 2011-2012 agricultural season.

Figure 11. Green LAI dynamics on three different fields, estimated using the SAFY model, and ground measurements during the 2011/2012 agricultural season.

Figure 12. Comparisons between measured and estimated (SAFY model) grain yields, for barley (a) and wheat (b).

Figure 13-a. Relationship between NDVI and leaf area index (LAI) measured in the test fields.

Figure 13-b. Comparison between measured and estimated LAI using equation (1), for the 2011-2012 agricultural season.

Figure 14. Cereal yield map produced by combining SPOT/HRV multi-temporal acquisitions and SAFY model data

Figure 15-a. Comparison between yield estimations computed using statistical remote sensing analysis, and yields determined by combining remotely sensed data with the SAFY model.

Figure 15-b. The difference map between the yields estimated by NDVI and by the SAFY model.

Table1. Satellite acquisition dates.

Table 2. Number and type of test fields for the two agricultural seasons.

Table 3. Satellite imagery used to produce the land use map.

Table 4. Ranges over which the three optimized SAFY parameters are considered.

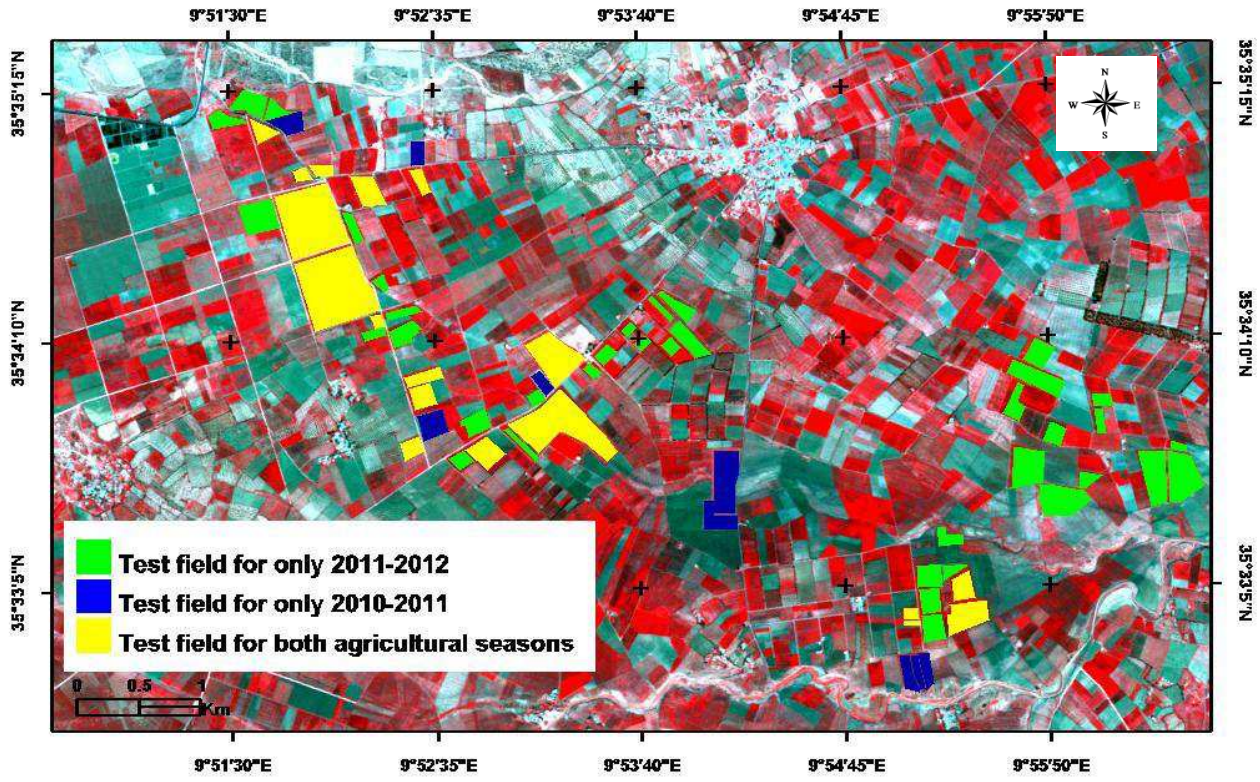


Figure 1. Illustration of the studied site.

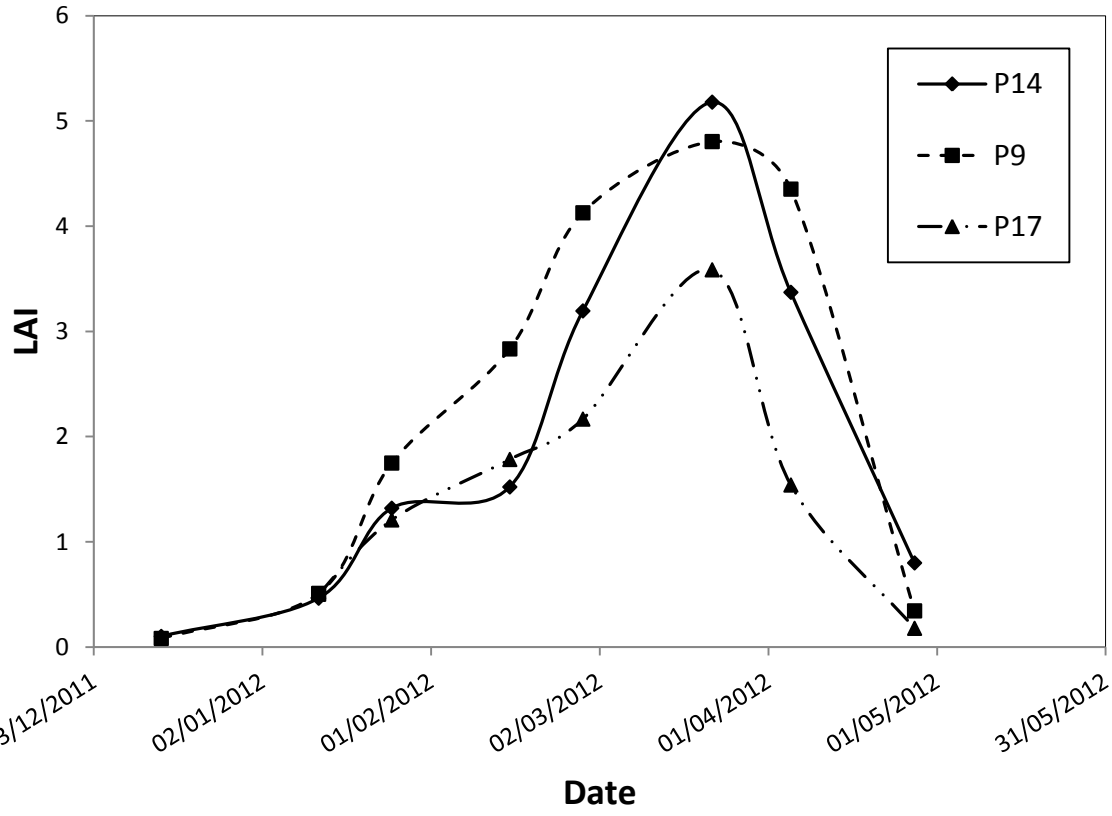
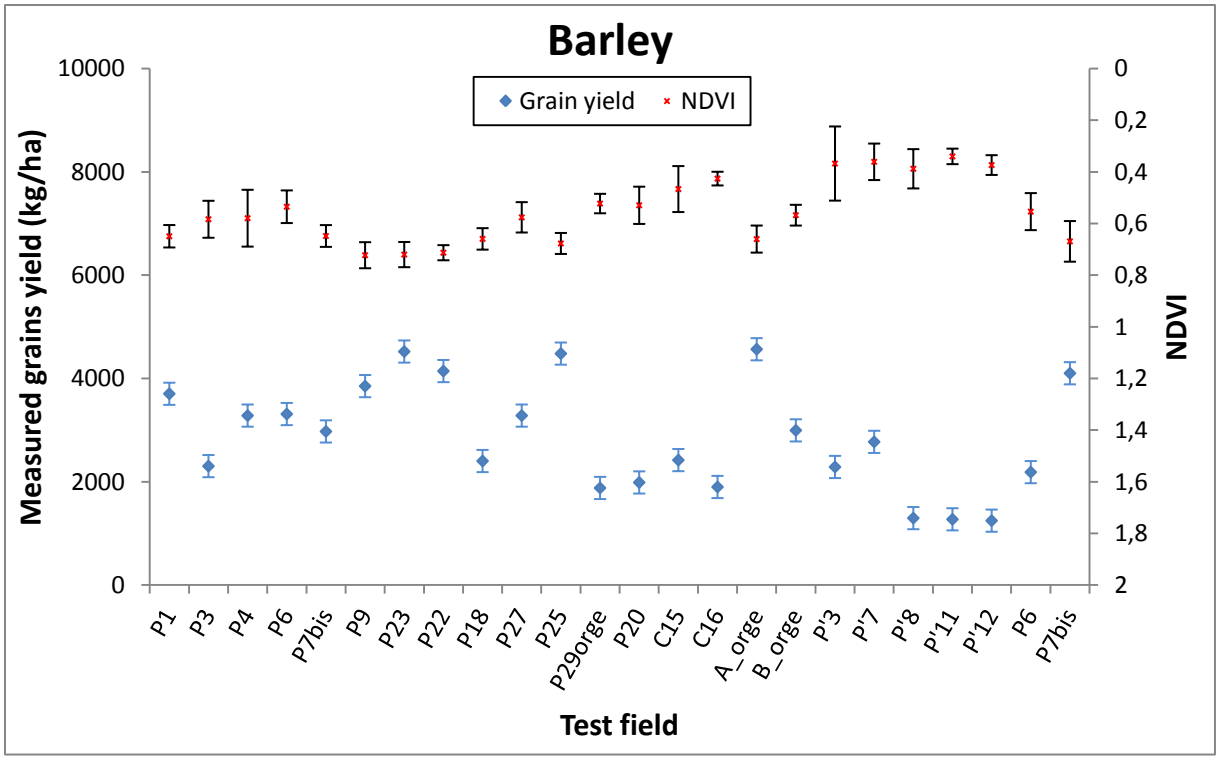


Figure 2-a. Illustration of the evolution of LAI measurements for three test fields: irrigated (P14 and P9) and rainfed (P17) fields, during the 2011/2012 agricultural season.



(a)

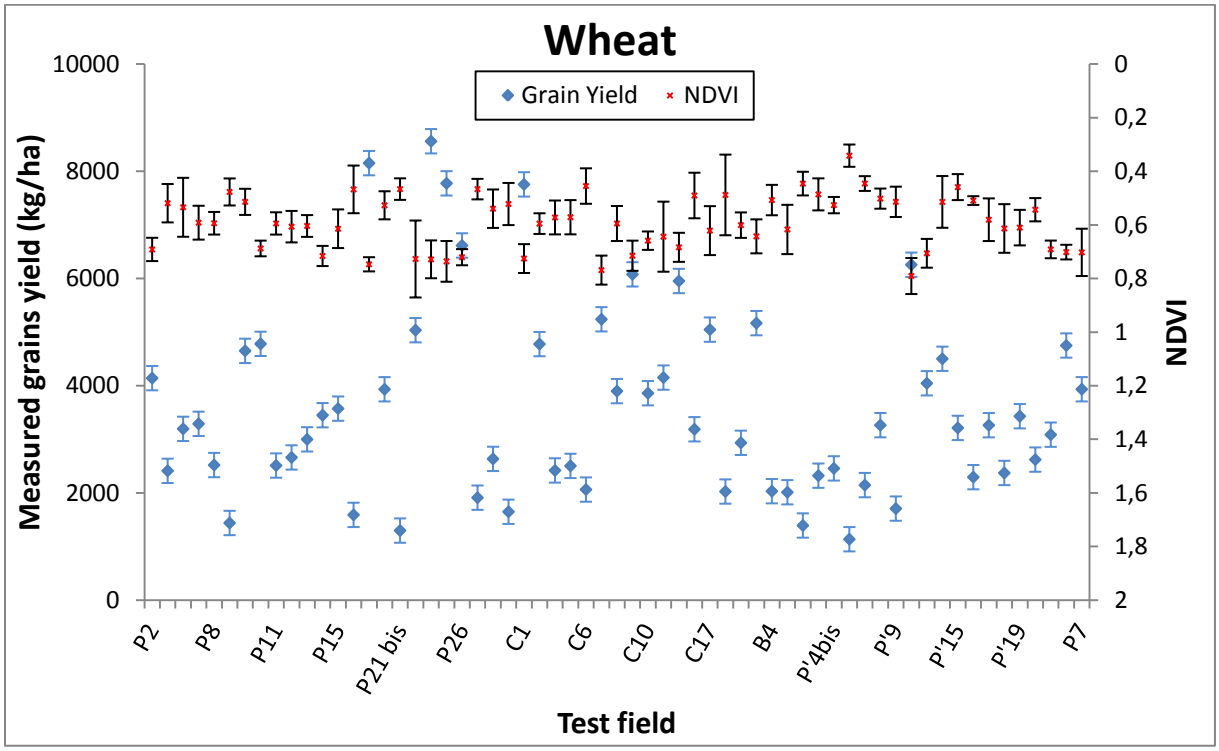


Figure 2-b. Comparisons between the error bars of measured grain yield and the NDVI extracted from satellite images for barley (a) and wheat (b).

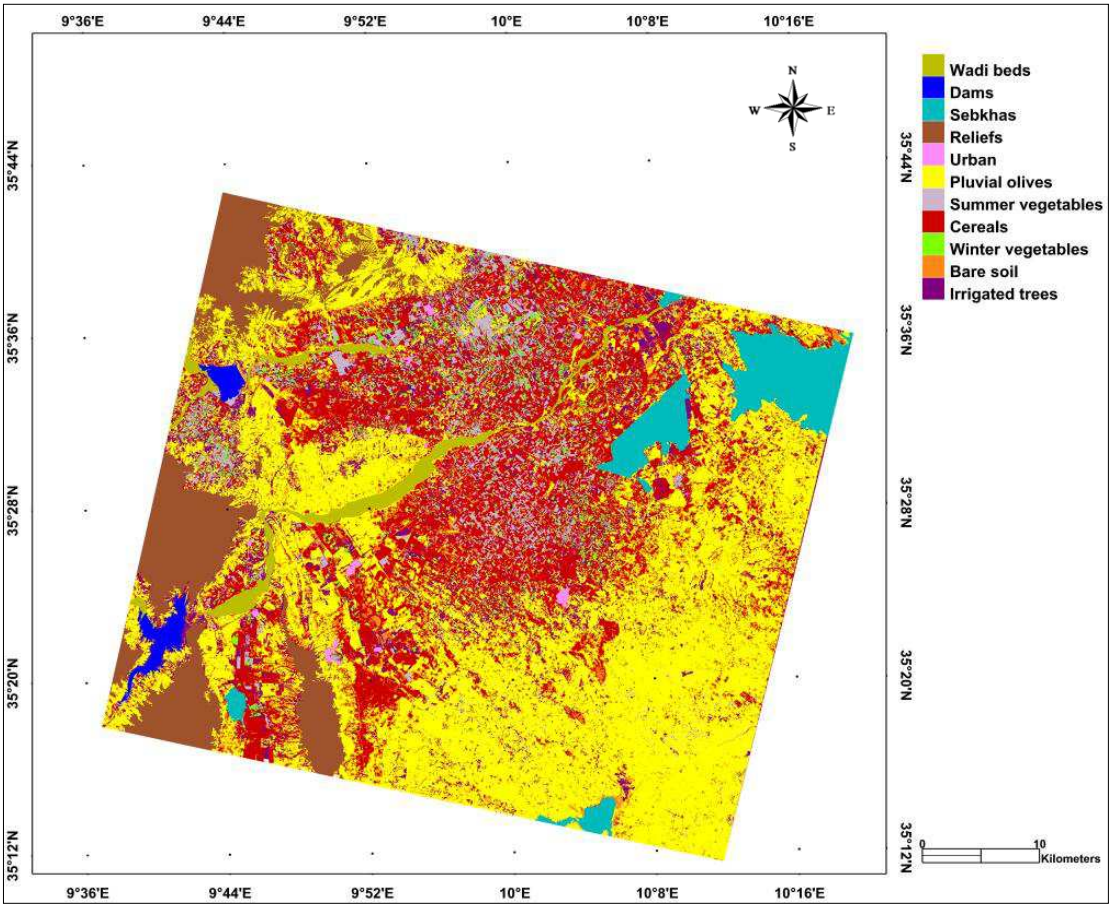
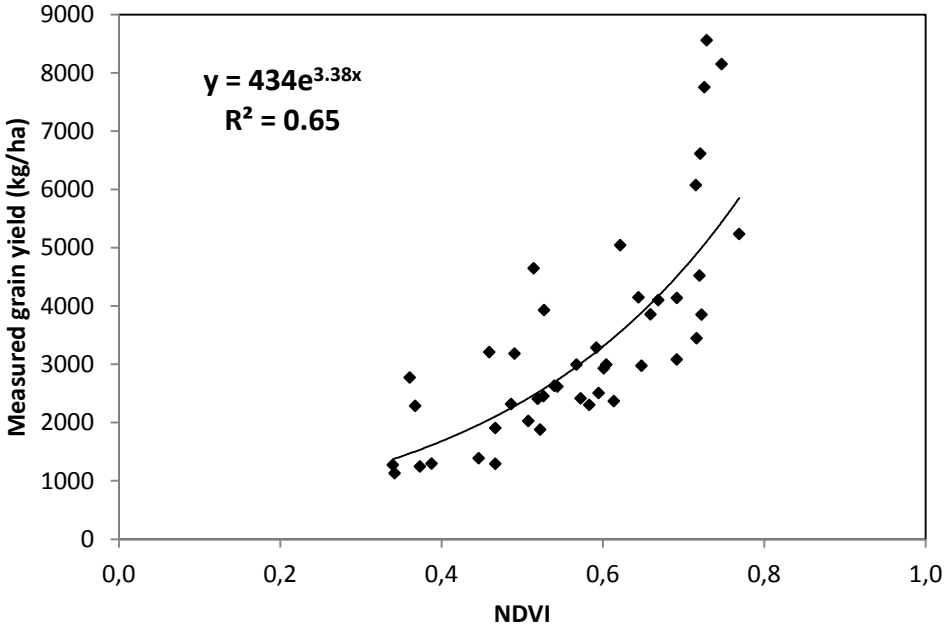
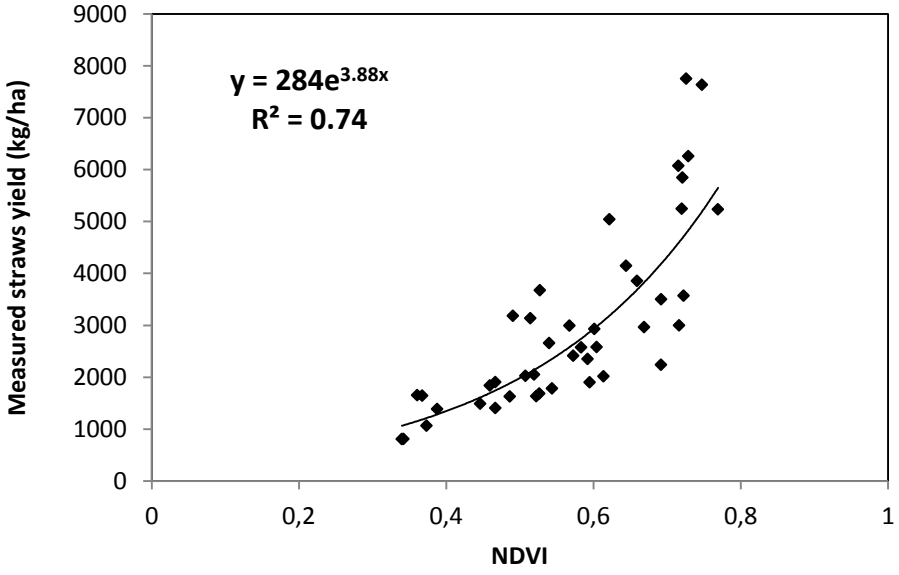


Figure 3. Land use map for the 2011-2012 agricultural season.

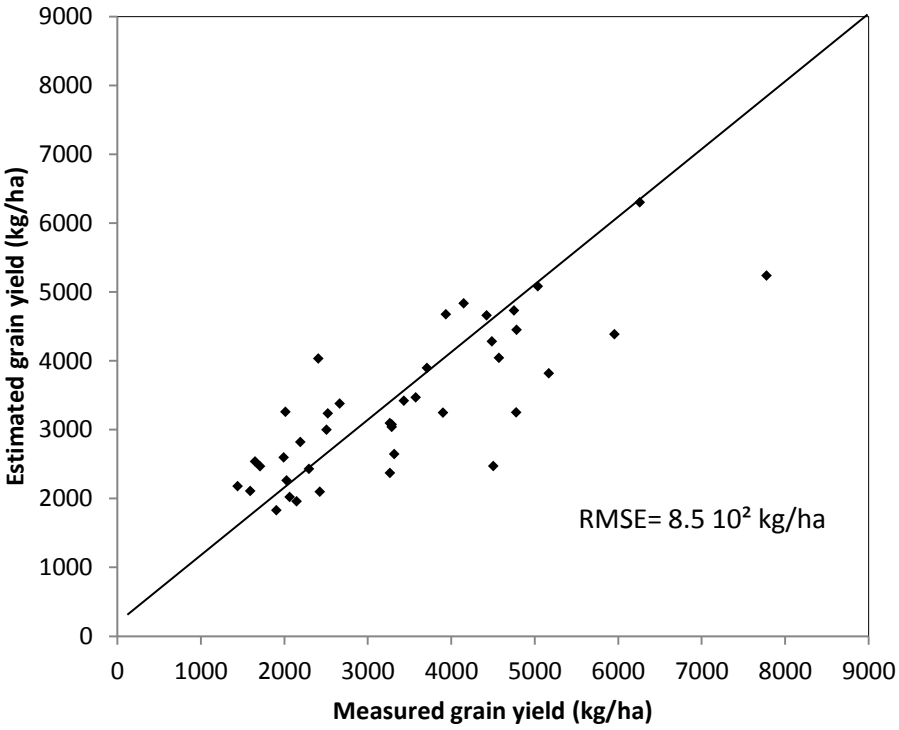


(a)

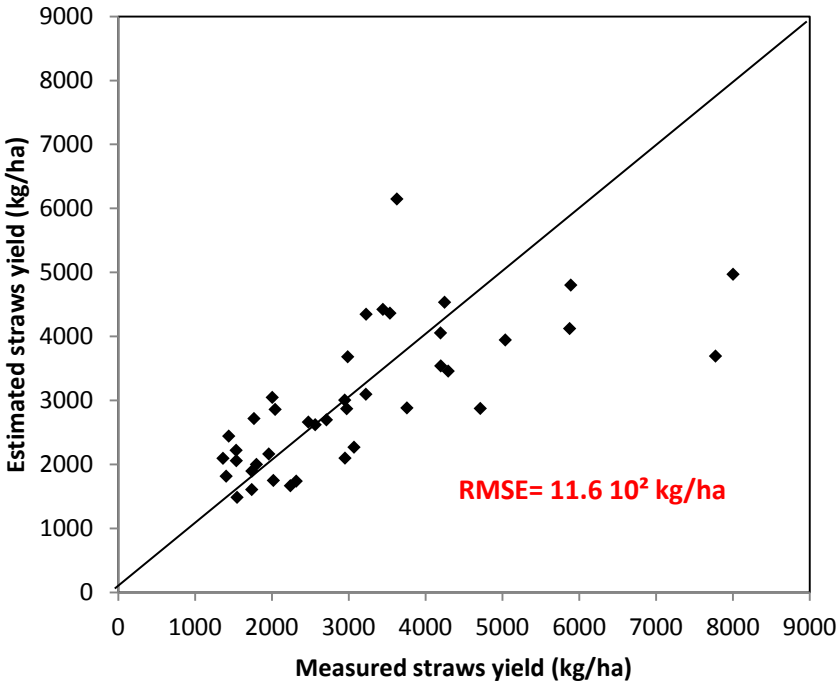


(b)

Figure 4. Relationship between mesured **cereal** yields (grain and straw) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012: (a) grain yields, (b) straw yields.

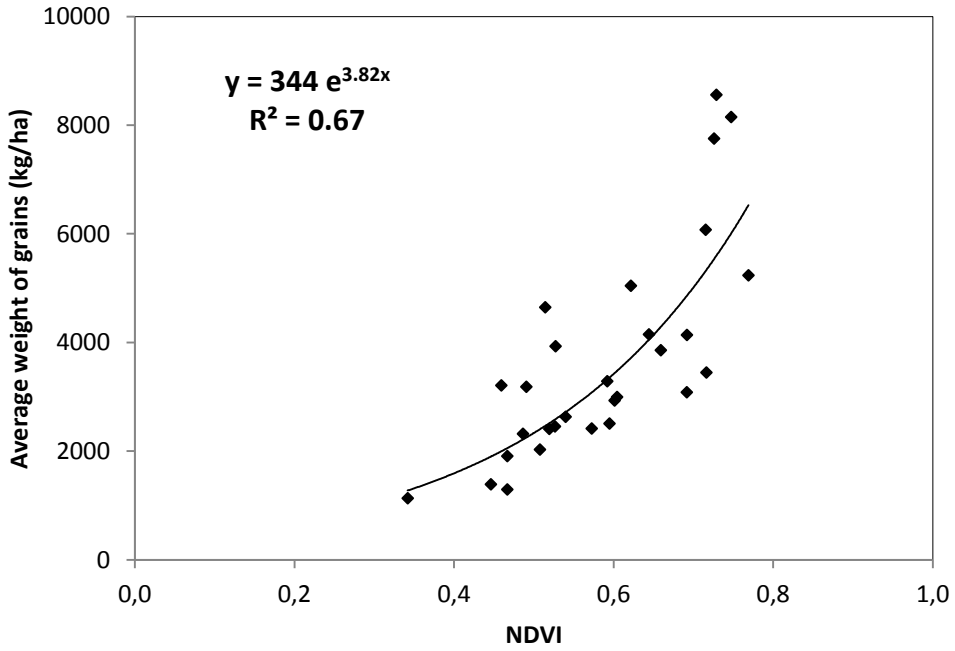


(a)

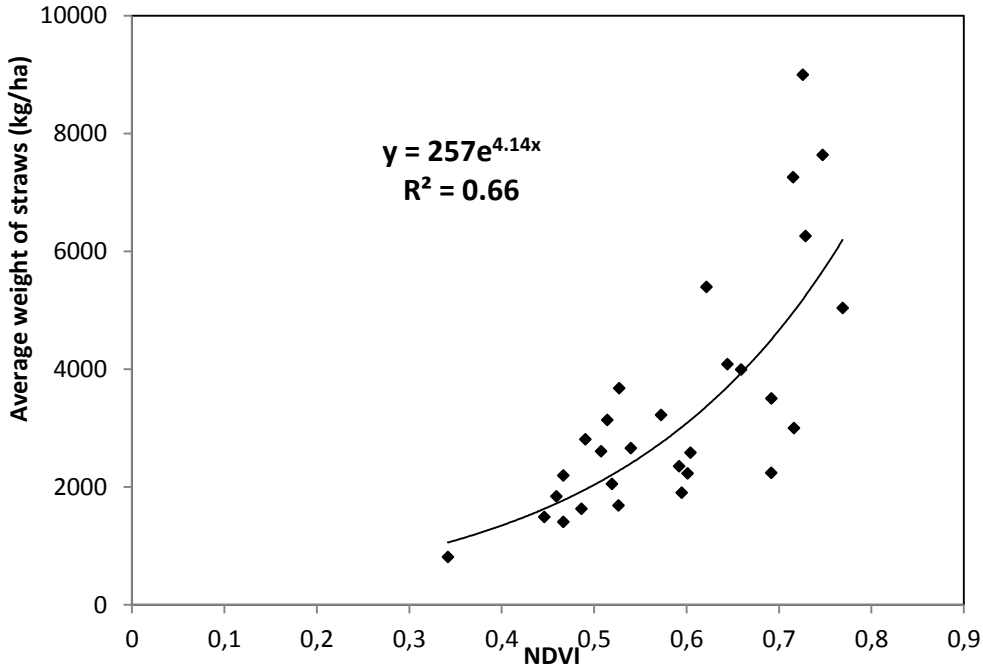


(b)

Figure 5. Comparison between measured and estimated **cereal** yields: (a) grain yields, b) straw yields.

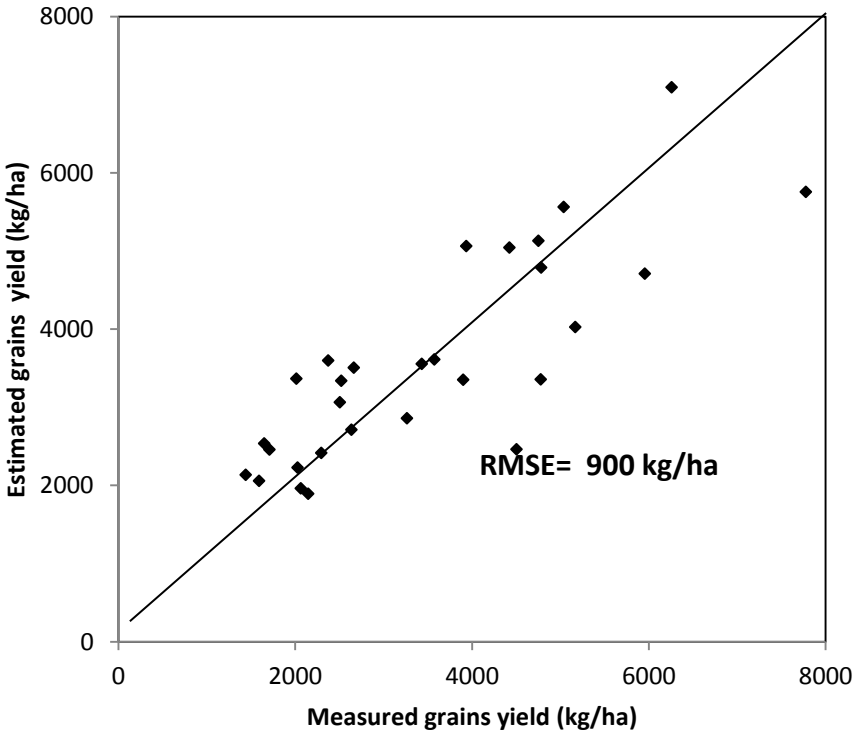


(a)

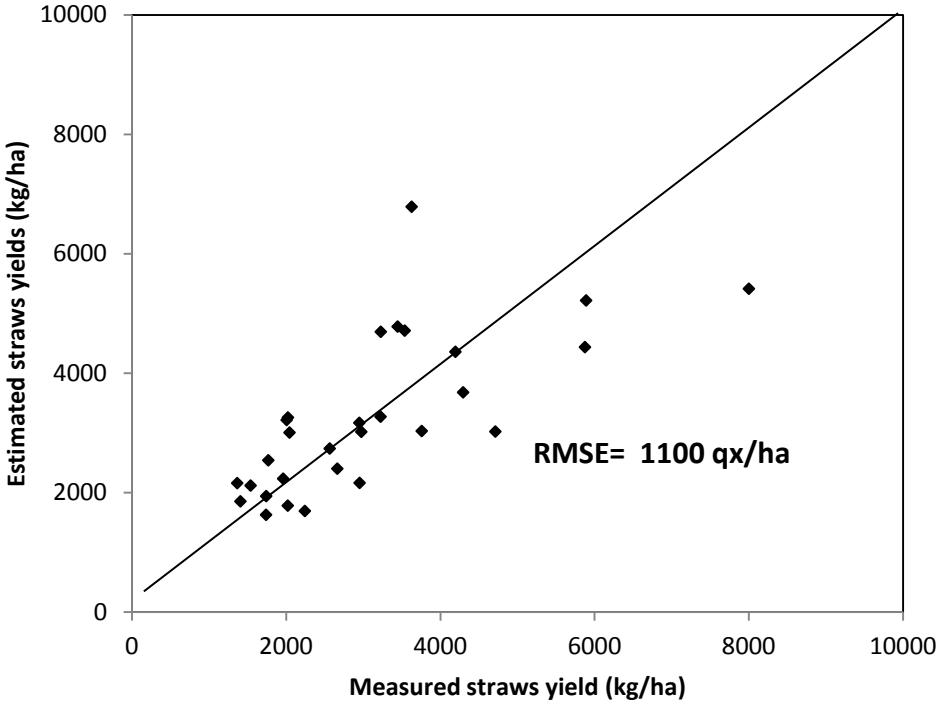


(b)

Figure 6. Relationship between mesured **whea**t yields (grain (a) and straw (b)) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012.

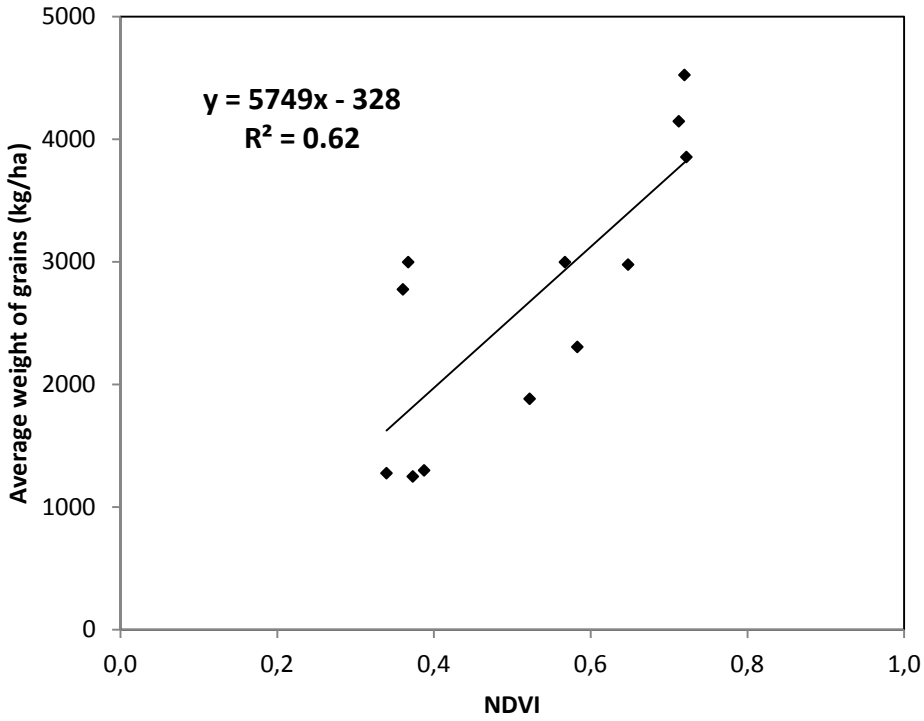


(a)

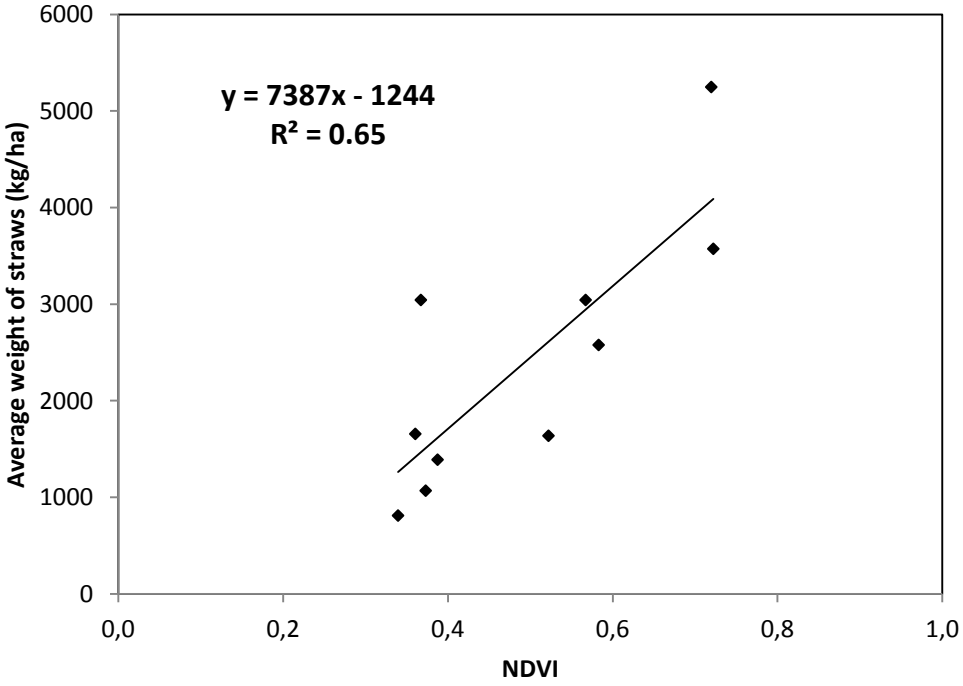


(b)

Figure 7. Comparison between measured and estimated wheat yields: (a) grain yields, (b) straw yields.



(a)



(b)

Figure 8. Relationship between measured barley yields (grain (a) and straw (b)) and the SPOT-NDVI index, on 17/03/2011 and 31/03/ 2012.